ISSN 1561-5359. Штучний інтелект, 2016, № 3

УДК 004.032.26

Е.Е. Марушко, А.А. Дудкин
Объединенный институт проблем информатики НАН Беларуси, Беларусь
ул. Сурганова, 6, г. Минск, 220012

ПОСТРОЕНИЕ АНСАМБЛЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ ТЕЛЕМЕТРИИ КОСМИЧЕСКИХ АППАРАТОВ

Y. Marushko, A. Doudkin
United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Belarus
Surganova,st., 6, Minsk, 220012

CONSTRUCTUIN OF ENSEMBLE OF NEURAL NETWORK FOR SPACECRAFT TELEMETRY MULTIVARIATE TIME SERIES FORECASTING

В работе предлагается подход к прогнозированию многомерных временных рядов телеметрической информации с использованием ансамблей нейронных сетей. Анализируются подходы к формированию ансамблей нейронных сетей и оценивается их точность. Изучается возможность обучения нейронных сетей для уменьшения ошибки прогнозирования временных рядов.
Ключевые слова: прогнозирование, нейронная сеть, ансамбль нейронных сетей, многомерный временной ряд, телеметрическая информация.

In this paper, we propose an approach to solving the problem of forecasting multivariate time series of telemetry data using ensemble neural networks. Approaches to the construction of neural network en-sembles are analyzed and accuracy of prediction is evaluated. It is studied a possibility of training the neural network ensembles for reducing errors of multivariate time series forecasting.
Keywords: Forecasting, artificial neural network, ensemble of neural networks, telemetry, multivari-ate time series.

Введение

Космическая телеметрия – это совокупность технологий, позволяющая производить дистанционный (удаленный) сбор информации о состоянии бортовых подсистем космических аппаратов. Их функционирование контролируется путем анализа показаний датчиков, которые распределены по модулям подсистем. Таким образом, телеметрия представляет собой многомерный временной ряд. Одной из задач анализа является прогнозирование данного ряда.

Постановка проблемы

Задача прогнозирования многомерного временного ряда в общем виде формулируется следующим образом [1, 2]: по известному текущему значению последовательности y(k) и некоторой предыстории y(k – 1), y(k – 2), …, y(k – m) дать оценку следующего значения $\hat{y}(k + 1)$. Каждый элемент последовательности y(k) представляет собой набор значений в момент k. Длина предыстории m, которая используется для прогнозирования, называется временным окном.

Широкое распространение для решения задач прогнозирования находит нейросетевой подход, так как он позволяет моделировать зависимости между данными в результате обучения на примерах. При этом эффективность использования нейронных сетей (neural network (NN)) зависит от архитектурных решений и методов обучения [1, 3], что приводит к необходимости многократного проведения

экспериментов. Имеется ряд примеров использования нейронных сетей в бортовых интеллектуальных системах поддержки принятия решений [4, 5].

В данной работе исследуется возможность кратковременного прогноза телеметрических параметров с использованием ансамблей нейронных сетей (ensemble of neural network, ENN) [6, 7], то есть набора NN, принимающего решения путем усреднения результатов работы отдельных NN.

В прогнозном анализе и машинном обучении имеет место дрейф, означающий, что статистические свойства целевой переменной, которую модель пытается предсказать, меняются с течением времени непредсказуемым образом [8], что увеличивает ошибку прогнозирования. Отсюда эффективность нейросетевой модели прогнозирования может быть повышена при помощи методов инкрементного обучения ENN [6, 8, 9].

Построение ансамблей для обработки телеметрии

Задача прогнозирования телеметрических данных решалась для трех подсистем белорусского космического аппарата. Это система энергоснабжения (power system, PS), корректирующая двигательная установка (corrective propulsion system, CPS) и целевая аппаратура (target equipment, TE). Соответственно, для обработки телеметрической информации (ТМИ) разработаны три ENN (рис. 1).

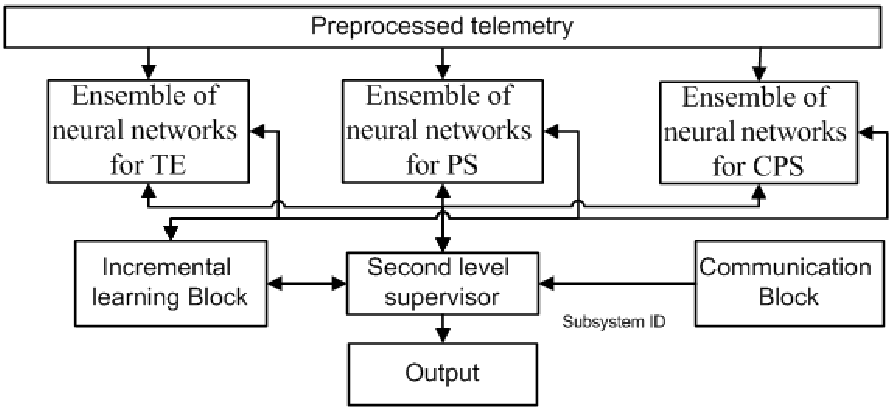


Рис. 1. Общая схема организации обработки с применением ENN

На вход ENN поступает предварительно обработанная ТМИ и идентификатор требуемой подсистемы, который с помощью коммуникационного блока передается в супервизор ансамблей. Супервизор формирует сигнал выбора ENN для конкретной подсистемы, получает от выбранного ансамбля сформированный прогноз, и инициирует процедуру дообучения. Блок инкрементного дообучения отвечает за подготовку набора обучающих данных и обучение новых элементов ENN.

В качестве базового элемента ENN используется многослойный персептрон с одним скрытым слоем с нелинейной функцией активации гиперболический тангенс.

Размер входного слоя N_I для NN одного ENN определяется как произведение числа датчиков подсистемы, для которой формируется ENN, и временного окна, по которому делается прогноз.

Размер скрытого слоя N_H для NN одного ENN определятся при проведении эксперимента с использованием процедуры поиска субоптимального размера скрытого слоя одиночной NN.

Размер выходного слоя N_O для NN одного ENN определяется числом датчиков подсистемы, для которой формируется ENN.

Обучение одиночных NN осуществляется алгоритмом RPROP [10].

Для формирования выходного значения ENN используются следующие подходы:
1) выходное значение ENN формируется как сумма выходов отдельных NN. Для случая с одним выходным нейроном вычисляется по формуле:

$$y = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (1)$$

где n – число моделей NN, y_i – выход i-й NN;
2) выходное значение ENN формируется как взвешенная сумма выходов отдельных NN. Для случая с одним выходным нейроном вычисляется по формуле:

$$y = \sum_{i=1}^{n} y_i \cdot w_i \qquad (2)$$

где n – число моделей, y_i – выход i-й NN, w_i – вес i-й NN, который формируется по формуле:

$$w_i = \frac{mse_i}{\sum_{i=1}^{n} mse_i}, \qquad (3)$$

где mse_i – MSE-ошибка i-й NN на валидационном наборе;
3) выходное значение ENN формируется как взвешенная сумма выходов отдельных NN (формулы 2 и 3), при этом взвешивание повторяется через определенный интервал обработанных наборов (временных отсчетов) с оценкой по этому набору (динамически взвешиваемый ENN).

Как видно, средневзвешенная сумма является частным случаем взвешенного суммирования при весовых коэффициентах равных 1.

Данный подход проиллюстрирован на рис. 2.

Понятие *дрейф значений* относится к изменению значения определения с течением времени и, следовательно, изменению в распределении данного значения. Среда, из которой эти значения получены, не является стационарной. Сдвиг в вероятности может указывать на то, что определения событий также могут изменяться.

Общим знаменателем в алгоритмах детектирования дрейфа выступает ансамбль экспертов, которые постепенно обучаются (без доступа к предыдущим данным) на входных данных, в сочетании с некоторой формой взвешенного голосования для получения финального решения [8, 9].

Инкрементное обучение ENN подразумевает оценку точности всех моделей и их ранжирование по точности на каждой итерации прогнозирования. При повышении ошибки ENN детектируется дрейф целевой переменной и производится добавление в ансамбль нового элемента, обученного на релевантных данных. При таком подходе сохраняется модель, заложенная при начальном обучении, и вносятся новые параметры без проблемы «забывания». Таким образом, реализуется дообучение ENN.

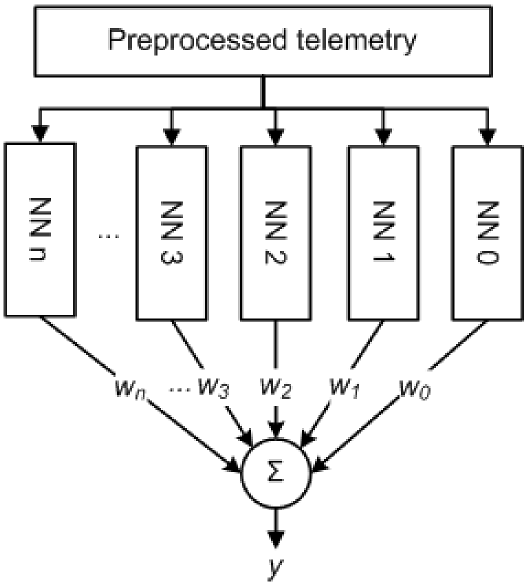


Рис. 2. Выходное значение ENN формируется как
взвешенная сумма выходов отдельных NN

Алгоритм инкрементного обучения ENN включает следующие шаги.

1. Производится оценка точности ENN путем сравнения ошибки на предыдущем шаге функционирования и текущем.
2. Если ошибка не изменилась, либо изменилась и находится в заранее заданном диапазоне, то осуществляется завершение работы.
3. Иначе производится формирование набора обучающих данных, который включает все накопленные данные с последнего дообучения.
4. Производится формирование и обучение новой NN.
5. Добавление сформированной NN в ансамбль.
6. Для всех NN ансамбля производится пересчет весовых коэффициентов по формуле (3).

Таким образом, двухуровневая модель ENN для прогнозирования телеметрии подсистем космического аппарата, представленная на рис. 1, формируется и работает по следующему алгоритму.

1. Указание размеров N_{PS} ENN_{PS} для обработки ТМИ PS, N_{CPS} ENN_{CPS} для обработки ТМИ CPS, N_{TE} ENN_{TE} для обработки ТМИ TE.
2. Обучение N_{PS} одиночных NN на наборе ТМИ PS.
3. Обучение N_{CPS} одиночных NN на наборе ТМИ CPS.
4. Обучение N_{TE} одиночных NN на наборе ТМИ TE.
5. Расчёт по формуле (3) весовых коэффициентов для одиночных NN каждого ENN_{PS}, ENN_{CPS}, ENN_{TE}.
6. Получение ТМ данных для обработки.
7. Определение обрабатываемой подсистемы (PS, CPS, TE).
8. Передача ТМ данных в соответствующий ENN.
9. Формирование результата по формуле (1) или (2).
10. Выполнение инкрементного дообучения выбранного ENN.
11. Если есть данные для обработки повторить п. 6.
12. Завершение работы.

Организация эксперимента для исследования эффективности анасамблей

В табл. 1 представлены параметры наборов данных, на которых проводился эксперимент.

Таблица 1. Данные телеметрической информации CPS

Название набора	*Время дискретизации, с.*	*Размерность временного ряда*	*Количество отсчетов*
Dt_set_s01	1	24	57501
Dt_set_s05	0.5	24	12245
Dt_set_s1	0.1	24	6613

Каждый из наборов представляет собой перечень телеметрической информации, формируемой датчиками CPS белорусского космического аппарата и характеризующей функционирование данной подсистемы: температурные параметры и уровни давления блока подачи ксенона, электрические параметры регулятора расхода, анода и катода двигателей. Значения, выдаваемые датчиками контролируемых параметров, зависят от режима работы CPS, который в свою очередь задается командами управления, передаваемыми с Земли. Значения датчиков CPS коррелируют между собой, что позволяет рассчитывать на удовлетворительные оценки прогноза.

Так как обучение выполняется с учителем, то необходимо сформировать обучающую выборку из пар «входной вектор, выходной вектор». Формирование пары обучающих примеров осуществляется по принципу скользящего окна [2]: т.е. берется некоторый отрезок временного ряда и из него выделяется несколько наблюдений, которые и будут представлять собой входной вектор. Значением желаемого выхода в обучающем примере будет следующее по порядку наблюдение. Затем окно сдвигается на одну позицию в направлении возрастания времени, и процесс формирования следующей пары обучающей выборки повторяется.

Таким образом, если размерность временного ряда данных равна N, а размер окна – W, то NN должна принимать на вход образец размером $N \times W$. Так, для $W = 20$ набор Dt_set_s1 преобразуется во входной набор размером 480× 6589 и целевой набор размером 28× 6589.

При подготовке генеральной совокупности телеметрической информации датчиков космического аппарата выполняется ресемплирование и масштабирование.

Ресемплирование выполняется для преобразования исходных данных, представляющих собой последовательность временных отметок важных событий, в форму с фиксированным временем дискретизации. Масштабирование необходимо, чтобы привести данные в допустимый диапазон [-1, 1]. Выходы сети также масштабируются.

Целью эксперимента является определение влияния характеристик одиночной NN и архитектур ENN на результативность прогноза, при окне прогнозирования равном 20 отсчетам.

Для оценки качества обученных NN и ENN, а также для сравнения различных архитектур ENN, используются следующие величины:

- средний квадрат ошибки (mean square error, MSE):

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m} e_i^2 \tag{4}$$

- средняя абсолютная ошибка (mean absolute error, MAE):

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m}|e_i|. \quad (5)$$

где $e_i = y_i - t_i$, y_i и t_i – полученный и желаемый сигналы на выходе i-го нейрона выходного слоя соответственно, m – размер выходного слоя NN.

В данном эксперименте входной набор разделялся в соотношении 9:1 на общую обучающую выборку и итоговую тестовую выборку. Общая обучающая выборка разделялась на валидационную (15 %), тестовую (15 %) и обучающую (70 %) выборки случайным образом, которые использовались для обучения, оценки и поиска лучшей архитектуры NN.

Итоговая тестовая выборка использовалась для вычисления конечных оценок полученных NN.

Оценка размера скрытого слоя NN

Оценивалась возможность поиска субоптимального размера скрытого слоя одиночной NN по следующему алгоритму.

1. Определение интервала поиска размера скрытого слоя.
2. Обучение 10 NN с размером скрытого слоя, равным текущему на интервале, формирование из 10 NN взвешенного ансамбля.
3. Оценка точности ENN.
4. Пока не достигнут конец интервала поиска, происходит переход к следующему элементу интервала.
5. Выбор ENN с наименьшей MSE на интервале поиска, размер скрытого слоя элемента ENN будет подходящим решением.

В табл. 2 приведены размеры скрытого слоя для наборов данных CPS.

Стоит отметить, что данный алгоритм значительно увеличивает время формирования архитектуры NN.

Таблица 2. Оценка размера скрытого слоя NN

Набор данных	*Размер скрытого слоя*
Dt_set_s1	28
Dt_set_s05	33
Dt_set_s01	70

Анализ подходов к формированию выходного значения ENN

На рис. 3 представлен график MSE в зависимости от размера скрытого слоя базовой NN.

В табл. 3 приведены оценки различных архитектур на тестовом наборе Dt_set_s1 с размером скрытого слоя одиночной нейронной сети согласно табл. 2.

Из данных в табл. 3 видно, что наименьшей ошибкой обладает динамически взвешиваемый ENN, при этом разница в оцениваемых параметрах между динамически и однократно взвешенными ENN очень мала.

Таблица 3. Оценка различных архитектур на тестовом наборе Dt_set_s1

Архитектура	*MSE, 10^{-4}*	*MAE, 10^{-3}*
Одиночная NN	3.51	0.115
ENN	3.66	0.114
Взвешенный ENN	2.76	9.65
Динамически взвешиваемый ENN, шаг = 10	2.75	9.59

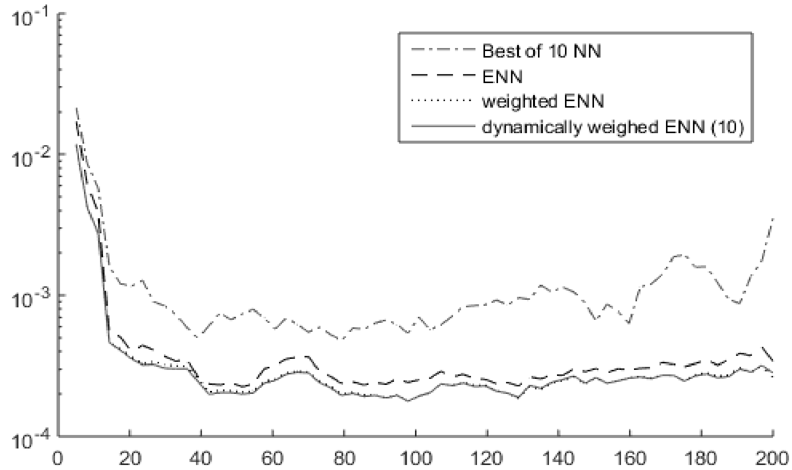


Рис. 3. Оценка различных архитектур в зависимости от размера скрытого слоя базовой NN

Оценка динамически взвешиваемого ENN

Проведена оценка шага взвешивания для динамически взвешиваемого ENN. На рис. 4 представлен график MSE для различных ENN.

Процедура оценки состоит из следующих шагов.

1. Обучение ENN с субоптимальным размером скрытого слоя.
2. Определение интервала поиска шага взвешивания.
3. Оценка MSE ENN, взвешенного ENN, динамически взвешенного ENN с фиксированным шагом, динамически взвешенного ENN с шагом, равным текущему значению из интервала.
4. Пока не конец интервала поиска, перейти к следующему элементу интервала.

Видно, что наименьшей ошибкой обладает динамически взвешиваемый ENN с малым шагом (менее 10 временных отсчетов).

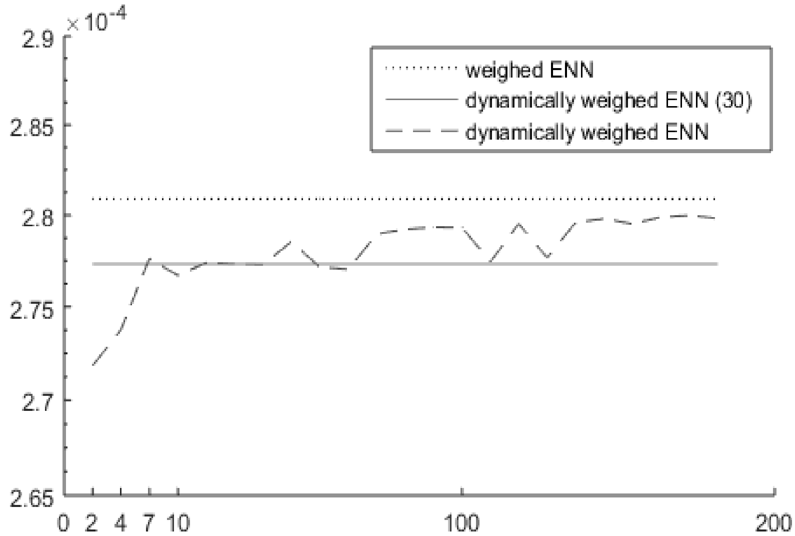


Рис. 4. Оценка динамически взвешиваемого ENN в зависимости от шага взвешивания

Оценка времени обучений нейронных сетей на различных вычислительных архитектурах

Целью эксперимента является определение различия в скорости выполнения обучения одиночной нейронной сети на центральном процессоре (ЦП) и графическом процессоре (ГП).

В качестве аппаратных средств использованы: ЦП Intel Core i5 4200H, 2 процессорных ядра, 4 потока выполнения, частота 2 800 МГц, Turbo-частота 3 400 МГц; ГП NVIDIA GeForce GTX 860M, частота 1029 МГц, 640 ядер CUDA.

В качестве экспериментальной платформы выбран набор средств Matlab. В эксперименте использовалась следующие параметры при обучении нейронной сети:
- архитектура – многослойная нейронная сеть с одним скрытым слоем;
- размер скрытого слоя в диапазоне [4, 56] нейронов;
- обучение методом RPROP;
- обучающая выборка 6613 образцов;
- размер входного слоя 480;
- размер входного слоя 1.

Процедура эксперимента: 1) Задать интервал размеров скрытого слоя. 2) Выбрать аппаратное средство (ЦП, ГП). 3) Выбрать значение из интервала. 4) Обучить 10 одиночных нейронных сетей. 5) Оценить среднее время выполнения. 6) Если не все значения из интервала использованы, то перейти к п. 3. 7) Если не все аппаратные средства использованы, то перейти к п.2. Результаты представлены в таблице 4.

Как видно из таблицы, в среднем наблюдается 5-ти кратное уменьшение времени обучения одной нейронной сети при использовании графического процессора с поддержкой технологии CUDA. Таким образом, применение ГП для ускорения обучения нейронных сетей позволяет значительно ускорить процедуру.

Таблица 4. Оценка времени обучений нейронных сетей на различных вычислительных архитектурах

Размер скрытого слоя	*Matlab CPU, с.*	*Matlab GPU(CUDA), с.*	*Ускорение*
4	93,9962	17,2422	5,4515
12	109,2202	19,5071	5,5990
20	86,2531	18,4156	4,6837
28	105,0203	20,0411	5,2402
36	113,4552	18,902	6,0023
44	90,6871	19,8095	4,5780
52	95,4515	16,4389	5,8064
56	107,1488	20,8944	5,1281
Среднее			5,3112
Медиана			5,3459
Минимум			4,5780

Заключение

Использование ENN значительно понижает ошибку при прогнозировании по сравнению с одиночной NN. Лучшей точностью обладает динамически взвешиваемый ENN с малым шагом взвешивания (менее 10 временных отсчетов), при этом разница в оцениваемых параметрах между динамически взвешиваемым и однократно взвешенным ENN незначительна. Наименьшая среднеквадратическая ошибка при анализе телеметрической информации, формируемой датчиками CPS, достигается при использовании динамически взвешиваемого ENN с шагом взвешивания 10 и равна $2{,}75\times10^{-4}$. Тестирование показало, что выполнение обучения одиночной нейронной сети на ГП с поддержкой технологии CUDA дает 5-ти кратное уменьшение времени обучения одной нейронной сети, по сравнению с ЦП, для указанного оборудования.

Литература

1. Short-Term Load and Wind Power Forecasting Using Neural Network-Based Prediction Intervals / H. Quan [et al.] // IEEE Transactions on Neural Networks and Learning Systems. – Short-Term Load and Wind Power Forecasting Using Neural Network-Based Prediction Intervals. – 2013. – Vol. 25, Iss. 2. – ISSN: 2162-237X. – P. 303 – 315.
2. Lysayak A.S. Prediction of multidimensional time series / A.S. Lysayak, B.Y. Ryabko // Vestnik SibGUTI. – 2014. – № 4. – P.75–88. (in Rusian)
3. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir / M. Valipour [et al.] // Journal of Hydrology 476 (2013). – P. 433–441.
4. Khachumov V.M. Review of Standards and the concept of monitoring, control and diagnostics of the spacecraft tools building / V.M. Khachumov [et al.] // – Software Systems: Theory and Applications. – №3(26), Volume 6. – 2015. – P.21-43. (in Russian)
5. Emelyanov Yu.G.: Neural orientation angles and distance of the spacecraft sensor control system / Yu.G. Emelyanov, K.A. Konstantinov, S.V. Pogodin, etc. // Software Systems: Theory and Applications, 2010. – № 1 (1). – P. 45-59. (in Russian)
6. Marushko Y. Using Ensembles of Neural Networks with Different Scales of Input Data for the Analysis of Telemetry Data / Y. Marushko // Proc. of the XV Intern. PhD Workshop OWD 2013 (Wisla, 19–22 Oct. 2013). – Gliwice: Silesian University of Technology, 2013. – P. 386–391.
7. Neural network ensemble operators for time series forecasting / N. Kourentzes [et al.] // Expert Systems with Applications. – July 2014. – Vol.41, Iss. 9. – ISSN: 0957-4174. – P. 4235–4244.
8. Elwell R. Incremental Learning of Variable Rate Concept Drift / R. Elwell, R. Polikar // MCS. – Vol. 5519 of Lecture Notes in Computer Science. – 2009. – P. 142–151.
9. Parikh D. An ensemble-based incremental learning approach to data fusion / D. Parikh, R. Polikar // IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics. – 2007. – Vol. 37. – № 2. – P. 437–450.
10. Riedmiller M. A direct adaptive method for faster back propagation learning: The RPROP algorithm / M. Riedmiller, H. Braun // In Proceedings of the IEEE International Conference on Neural Networks (ICNN), San Francisco. – 1993. – P. 586–591.

RESUME

Y. Marushko, A. Doudkin

Construction of Ensemble of Neural Network for Spacecraft Telemetry Multivariate Time Series Forecasting

In this paper, we propose an approach to solving the problem of forecasting multivariate time series of telemetry data using ensemble neural networks. Approaches to the construction of neural network ensembles are analyzed and accuracy of prediction is evaluated. It is studied a possibility of training the neural network ensembles for reducing errors of multivariate time series forecasting. Experiments show that ensembles significantly reduce the forecasting error in comparison with single neural networks. Dynamically weighted ensemble with a small step of repetitive weighing (less than 10 samples) has the smallest error and the difference in the estimated parameters for weighted ensembles is very small. The minimum mean square error at short-term forecasting of telemetry data obtained by sensors of the correction propulsion system is equal to 2,75×10-4. Also testing has shown that neu-ral network training on a single GPU with CUDA technology enables 5-fold decrease in the time of training a single neural network in comparison with the CPU in the specified equipment.

Надійшла до редакції 12.09.2016